时序视域下图情领域重要国家的学术贡献研究*

■ 刘虹　李煜
南京大学信息管理学院　南京 210023

摘要:[**目的/意义**]揭示2002年以来图情领域重要国家在学术论文产出的总量、引证、研究主题、学术合作等维度上的学术贡献与特征演化,为图情领域的学科建设及相关研究者追踪研究主题与前沿提供参考。[**方法/过程**]以近15年WoS数据库收录的图情领域86种核心期刊为研究对象,应用文献计量学、科学知识图谱、计量模型等研究方法,运用TDA、Ucinet、VOSviewer等数据分析工具,从学术论文的总量特征、引证特征、主题特征三个角度对图情领域重要国家的学术贡献进行深入分析,并构建计量模型考察空间距离及学术能力对重要国家之间学术合作的影响机制。[**结果/结论**]美国在图情领域的科研实力强劲,中国的学术贡献增速最快,荷兰、芬兰、比利时三国则在归一化被引比率指标上表现突出;研究主题主要集中在图书馆、信息检索、互联网、知识管理、社交网络、计量学、用户研究、研究方法8个主题范畴,英国、西班牙、中国、美国的研究主题最为广泛;空间距离与学术能力都对国家间的学术合作具有显著影响。

关键词:图书情报　科研合作　信息计量　知识图谱　可视化
分类号:G250
DOI:10.13266/j.issn.0252-3116.2018.23.011

1 引言

高水平学术论文能够表征科研机构在国际视野下的学术水平与学术影响力[1],也是国家双一流建设的成果体现与必要条件。本文通过对时序视域下图情领域重要国家的学术贡献的研究,可以揭示图情领域不同国家在学术论文产出的总量、文献引证、研究主题等维度的学术能力与学术地位的阶段性特征。基于研究对象,可将国内有关图情领域研究成果划分为3类主题:①以图情领域特定主题为研究对象的学术论文研究脉络分析。例如施萧萧等[2]、李信等[3]分别以“大数据”“颠覆性创新”为研究主题,对WoS数据库收录的相关主题文献展开共词分析、聚类分析并绘制知识图谱,揭示了相关主题研究的主题结构、研究热点及前沿趋势。②以我国图情领域学术论文为研究对象的研究动态分析,包括对研究主题、知识基础、发展态势的研究和国际学术合作特征的研究。例如谭春辉等[4]以9种情报学CSSCI来源期刊为研究对象,从核心机构、核心作者与热点主题角度对“十二五”时期情报学的研究态势进行了细致研究;赵蓉英等[5]基于比较静态分析方法、童寿传等[6]基于参考文献、施引文献分别对我国图情领域的国际化合作与交流的特征演化展开深入研究。③以国家间的图情领域学术合作与发展动态为研究对象的研究。例如侯剑华等[7]应用多视角的共被引分析方法揭示了2009年以来国际图情领域的研究概况、主题结构及研究热点。

以上研究展现了特定时间段内图情领域研究的阶段性特征。纵观来看,随着数据密集型科学研究的兴起,图情领域的学术论文产出、研究主题、研究力量演化出了新的发展趋势。在考察各国在图情领域的学术贡献时,不仅需要考察各国的学术论文产出总量,还需综合分析各国的学术论文产出质量以及在学科前沿领域的成果数量和科研合作中的学术地位。本文搜集了2002年-2016年的图情领域SCI/SSCI所有来源期刊的论文的总量数据和引证数据,基于时序视域,综合采用文献计量学、科学知识图谱、计量模型等研究方法,从学术论文的总量特征、引证特征、主题特征三个角度,力图在更大的时间尺度和空间范围内对图情领域

* 本文系国家自然科学基金项目“融合范式视角下的链接分析理论集成框架及其实证研究”(项目编号:71273125)研究成果之一。
作者简介:刘虹(ORCID:0000-0003-4943-7855),博士研究生,E-mail:liuh.1990@163.com;李煜(ORCID:0000-0001-8385-9862),馆员,博士研究生。
收稿日期:2018-05-10　修回日期:2018-08-11　本文起止页码:87-96　本文责任编辑:杜杏叶

各国的学术贡献展开分析，并构建计量模型考察影响重要国家之间科研合作的关键因素。

2　学术贡献研究的框架设计及实现机制

2.1　框架设计

为了更加全面地分析科图情领域主要国家的学术贡献，本文构建学术贡献研究体系如图 1 所示：

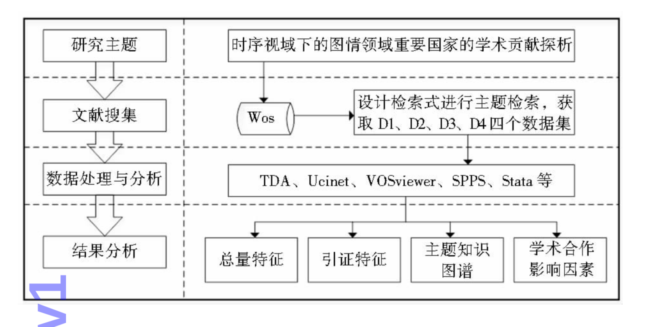


图 1　学术贡献研究的体系框架

首先，检索式构建与数据搜集。本文以 WoS 核心集（SCI、SSCI、A&HCI）中图情领域的 86 种核心期刊为研究对象，限定年限 2002 年 – 2016 年，论文类型为"Article"，以"SU = library and information science"为表达式进行高级检索，获取近 15 年的文献数据记为数据集 D1，共包括 154 个国家（地区）的 47 758 条文献数据；同时，以作者地址字段为准、不区分作者排序，即在同一篇文章中出现的多个国家不分排序先后、出现多次不累加，只要出现即认为产出一篇文章，截取发文量占比 1% 以上的 19 个重要国家的 39 647 条文献数据记为数据集 D2；获取 19 个重要国家在 2002 年 – 2016 年时期内历年的参考文献共 913 877 条，记为数据集 D3；使用 WoS 的"创建引文报告"功能逐项获取重要国家 t 年（1999 年 – 2013 年）所发表文献在 t + 3 年（2002 年 – 2016 年）的被所有 154 个国家（地区）的引用数据共 81 970 条，记为数据集 D4。检索时间为 2018 年 1 月 10 日 – 1 月 13 日。

其次，数据清洗。本文的数据清洗对象主要为国家字段和关键词字段。①国家字段清洗：同一国家不同地区字段的合并，如中国大陆地区（含香港、澳门）与台湾地区的数据合并，英格兰、苏格兰、北爱尔兰的数据合并。②关键词字段清洗：对比分析发现，WoS 题录中的作者关键词比系统关键词更能表征文章研究主题，表明图情领域学者文献格式的使用和学术素养都处于较高水平。为更好厘清学科领域的知识主题，需要对作者关键词进行规范化处理，具体操作包括合并同义关键词、拆分复合关键词、剔除无效关键词，如 WoS 中"信息素养"就包括 information literacy、information literacies、informational literacy、literacy：information 4 种表述方式，需对其进行规范化处理。最终，将数据集 D1 中的 72 086 个关键词精简为 44 406 个。

第三，高频词的选取与共现矩阵构建。选取发文量占比 1% 以上的 19 个国家作为重要国家；J. C. Donohue[8] 基于齐普夫第二定律提出确定高频关键词临界值的算法，本文引入该算法确定高频关键词数量（公式 1）。式中，I_n 是高频关键词数量，I_1 是文章中出现频次为 1 次的关键词数量。

$$I_n = \frac{\sqrt{8I_1 + 1} - 1}{2} \quad (1)$$

第四，数据分析。具体包括学术论文主题的可视化分析与学术合作的影响因素的实证分析。

2.2　研究方法与实现机制

本文以规范化处理后的 4 个数据集 D1、D2、D3、D4 为来源数据，从以下 4 个方面对图情领域重要国家的学术贡献展开可视化分析。

首先，基于数据集 D2、D3，采用时序分析方法，对比分析重要国家在图情领域的历年学术论文总量特征与被引用概况，分析各国的论文产出总体趋势。

其次，基于数据集 D4，借鉴 L. Bornmann 等[9] 提出的学术论文归一化被引比率（公式 2）计算得出图情领域重要国家的历年归一化被引比率，比较各国的学术论文在引证维度上的影响力大小。其中，$R_{a,t}$ 是 a 国在 t 年的归一化被引比率，$P_{a,t}$ 和 $C_{aj,t+3}$ 分别是 a 国在 t 年发表的文献数量和这些文章在 t + 3 年被 j 国引用的次数，M 是在 t 年所有发表文章的国家数量，N 是在 t + 3 年所有引用 t 年所发文章的国家数量。

$$R_{a,t} = \frac{\sum_{j=1}^{N} C_{aj,t+3} / \sum_{i=1}^{M} \sum_{j=1}^{N} C_{ij,t+3}}{P_{a,t} / \sum_{i=1}^{M} P_{i,t}} \quad (2)$$

第三，基于数据集 D1、D2，绘制图情领域的主题知识图谱。应用 Ucinet 中集成的 NetDraw 软件绘制不同时间窗口下的主题知识图谱以及国家 – 主题二模知识图谱，应用 VOSviewer 绘制不同时间窗口下的重要国家学术合作图谱，把握时序视域下图情领域研究主题的演化趋势，应用内容分析法总结归纳图情领域主题范畴及其在不同国家的分布。

第四，基于数据集 D2，探索重要国家之间学术合作的影响因素。基于空间距离与学术能力变量构建图情领域学术合作影响因素计量模型，应用 Stata 对模型进行实证研究。

刘虹，李煜. 时序视域下图情领域重要国家的学术贡献研究[J]. 图书情报工作,2018,62(23):87－96.

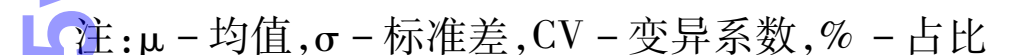
3 学术论文产出的总量特征分析

3.1 学术论文产出总量的时序分析

基于文献信息计量学的分析方法,各国图情领域的学术论文产出数量可以直观反映其在该领域的科研能力与学术地位[10](见表1、图2)。

表1 重要国家学术论文产出总量特征

排序	国家	μ	σ	CV	%	排序	国家	μ	σ	CV	%
1	美国	1255.87	135.14	0.11	41.86%	11	法国	52.13	21.81	0.42	1.74%
2	英国	299.27	33.96	0.11	9.98%	12	意大利	50.53	26.17	0.52	1.68%
3	中国	256.53	156.98	0.61	8.55%	13	比利时	47.13	13.34	0.28	1.57%
4	西班牙	172.53	100.02	0.58	5.75%	14	新加坡	42.20	13.54	0.32	1.41%
5	加拿大	165.67	38.82	0.23	5.52%	15	瑞典	41.53	17.54	0.42	1.38%
6	澳大利亚	125.87	52.99	0.42	4.20%	16	芬兰	41.40	15.37	0.37	1.38%
7	德国	103.53	42.77	0.41	3.45%	17	印度	40.73	16.74	0.41	1.36%
8	荷兰	89.53	35.31	0.39	2.98%	18	南非	36.33	17.79	0.49	1.21%
9	韩国	78.80	47.77	0.61	2.63%	19	丹麦	31.87	13.34	0.42	1.06%
10	巴西	68.53	48.43	0.71	2.28%						


注:μ－均值,σ－标准差,CV－变异系数,%－占比

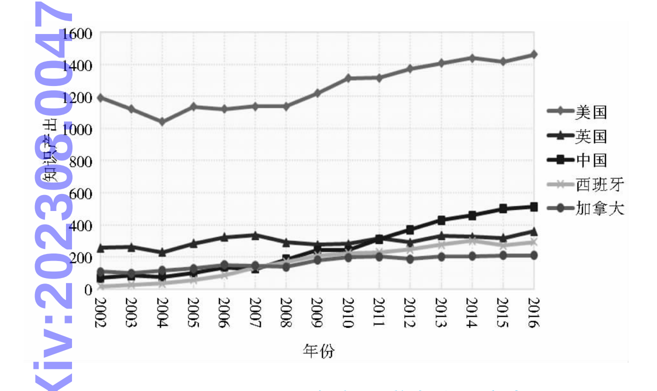


图2 2002－2016年各国学术论文产出

采用数据分析软件SPSS22.0对重要国家的学术论文产出进行曲线拟合发现,线性函数 $y=0.046x^2-181412.72$($R^2=0.98$)最符合重要国家的学术论文产出总量随时间演变的发展趋势。同时,美国、英国两国学者在WoS核心数据库上发表的文章总量分列第一、第二,其中美国在图情领域的学术论文产出遥遥领先于其他各国,占所有重要国家学术论文产出总量的41.86%。另外,中国和西班牙在图情领域的发文势头迅猛,中国自2004年学术论文产出开始迅速增加并于2012年赶超英国,西班牙也于2008年开始稳居学术论文产出总量第四名。加拿大、澳大利亚、德国、荷兰、韩国、巴西六国在图情领域的学术论文产出数量在位居前十。

3.2 基于学术论文产出的成分聚类

为了更好地反映随着时间演变各国学术论文产出的变动差异程度,本文引入变异系数(CV,标准差与平均值之比),变异系数数值越大表明在近15年内该国历年的学术论文产出变动越大。重要国家的学术论文产出均值和变异系数均值分别为158和0.41。本文借鉴波士顿矩阵思想,将学术论文产出均值和变异系数均值作为坐标轴构建二维矩阵,从而将19个国家划分为4个K成分(见图3)。

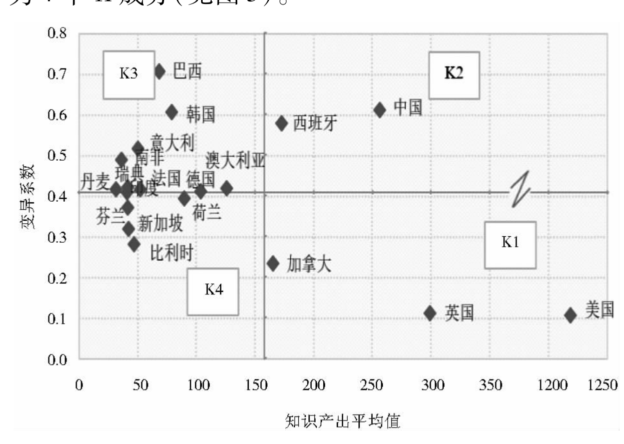


图3 学术论文产出K成分

在4个学术论文产出成分中,K1成分特征为学术论文产出占有率高,但增长率低于平均水平,由美国、英国、加拿大三国组成,它们是图情领域研究的传统科研强国;K2成分特征为学术论文产出占有率和增长率都高于平均水平,由中国和西班牙组成,它们是图情领域研究的新兴科研强国;K3成分特征是学术论文产出占有率较低,但增长率高于平均水平,由巴西、韩国、意大利、南非等十国组成,它们是图情领域研究的科研潜力国家;K4成分特征为学术论文产出占有率和增长率都低于平均水平但增长率仍保持上升势头,由芬兰、荷兰、新加坡、比利时四国组成,它们是图情领域研究的

科研增速放缓国家。

4 学术论文产出引证特征分析

引证特征视角下的学术论文产出分析能够客观评估各国在图情领域研究成果的被认可和被关注的程度，在一定程度上也能衡量其在国际知识交流与合作过程中的影响力[11]。本文采用被引量和归一化被引比率（Ratio of Citations，RoC）两个指标对重要国家的学术论文产出的引证特征展开时序分析。

基于数据集 D3 对重要国家在 WoS 核心集的年均被引量进行时序分析，基于数据集 D1、D2、D4，对重要国家在 WoS 核心集的年均归一化被引比率进行时序分析。其中，被引量指标体现的是各国的学术论文产出累积水平及其被认可与接受程度在历年的可见度，而归一化被引比率则体现了各国当年学术论文产出的被认可和接受程度，不受前期成果的影响。以美国为例，2002 年美国学术论文产出被引量即为 WoS 核心集上 2002 年收录的所有文章对美国 2002 年及以前年份发表的文章的引用次数；2002 年美国归一化被引比率则为美国 2002 年发文量占比与这些文章在 2005 年被所有国家引用次数占比的比值。

4.1 被引量的时序分析

重要国家学术论文产出的被引量分析结果见表 2 和图 4。

表 2 重要国家学术论文产出引证特征

排序	国家	μ	σ	%	排序	国家	μ	σ	%
1	美国	3.15	16 508.35	51.74%	11	比利时	0.09	556.62	1.47%
2	英国	0.61	3 065.12	10.01%	12	芬兰	0.08	540.36	1.38%
3	加拿大	0.49	2 704.23	8.03%	13	法国	0.08	477.67	1.26%
4	中国	0.37	3 538.72	6.14%	14	丹麦	0.06	376.24	1.07%
5	荷兰	0.25	1 669.51	4.07%	15	瑞典	0.06	471.44	1.03%
6	澳大利亚	0.21	1 346.94	3.44%	16	意大利	0.06	466.78	0.99%
7	德国	0.13	1 043.07	2.11%	17	印度	0.06	466.78	0.99%
8	西班牙	0.13	1 227.26	2.09%	18	南非	0.03	196.04	0.44%
9	韩国	0.12	983.51	1.91%	19	巴西	0.02	191.72	0.38%
10	新加坡	0.09	602.70	1.49%					

注：μ－均值，σ－标准差，%－占比

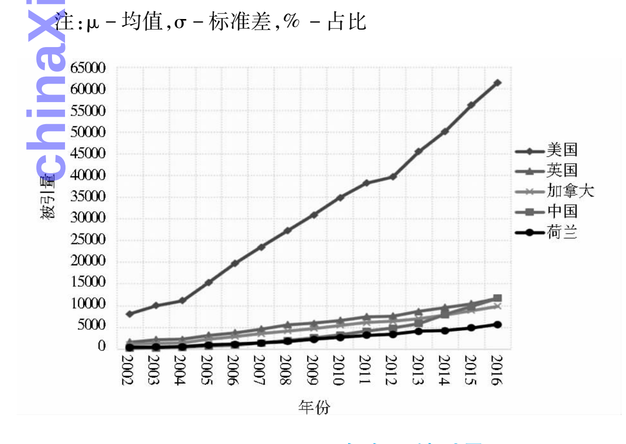


图 4 2002－2016 年各国被引量

美国的被引量增速远超其他各国且被引量均值占比为 51.74%，英国、加拿大、中国的被引量分列第二、第三、第四。结合图 5 的学术论文产出与被引量排名展开对比分析，美国、英国的学术论文产出总量和年均被引量在图情领域都稳执牛耳，这与两国政府、基金资助机构长久以来对图书情报领域的高度重视以及相关科研院所、科研人员在该领域学术成果的长期积累紧密相关。其次，加拿大与美国、英国同属 K1 科研强国队列，其学术论文产出的影响力也十分可观，在 19 个重要国家中名列第三。同时，中国在近 15 年内图情领域的学术论文产出和被引增速显著，总量分别位居第三、第四位，虽然其被引量排名略低但仍表现出了良好的发展趋势，科研实力强劲。另一方面，巴西、西班牙、意大利的被引量排名显著低于学术论文产出数量，三国需提高对学术论文产出质量的把控以提高其学术成果的被认可和接受程度；丹麦、芬兰、新加坡的学术论文产出排名较为靠后但其被引量排名则明显优于其学术论文产出排名，是小而精典范；澳大利亚、德国、韩国、瑞典等国学术论文产出和被引量排名则较为一致。

4.2 归一化被引比率的时序分析

根据公式 2，得出 1999 年－2013 年重要国家学术论文产出的历年的归一化被引比率（RoC），进一步计算 RoC 均值（见表 3）并按该指标将这些国家划分为三个期望子群，各子群内部每个国家的归一化被引比率随时序的演化过程见图 6。

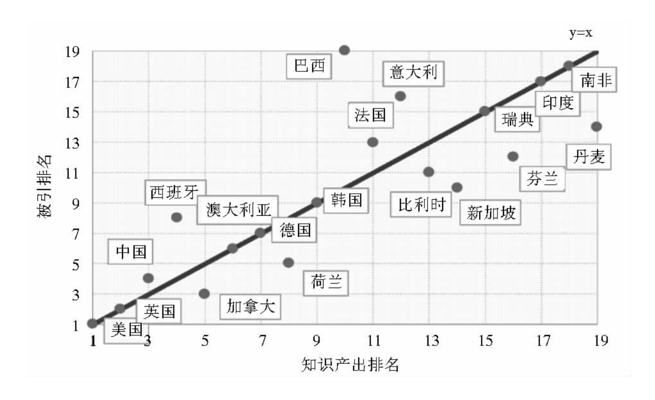
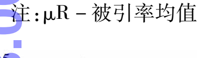


图 5　重要国家学术论文产出与被引排名

(1)超出期望(RoC≥1.2):11 个国家的 RoC 均值超出期望,排名前三的国家分别为荷兰、芬兰、比利时,它们的学术论文产出总量虽然不多,但是学术论文产出水平较高,在国际范围内得到了大量引用;美国和加拿大的 RoC 均值也高于 1.2,表明两国在保证学术论文产出高速发展的同时其学术论文产出质量也同样出类拔萃,以此趋势可预见两国将在图情领域继续保持科研强国和科研大国地位;同时,中国的 RoC 均值高达 1.53,无论是在学术论文产出总量还是产出水平来看,中国都表现出了强劲的科研实力。

表 3　重要国家归一化被引比率均值

排名	国家	μR	排名	国家	μR	排名	国家	μR	排名	国家	μR
1	荷兰	2.01	6	韩国	1.46	11	美国	1.21	16	南非	0.75
2	芬兰	1.65	7	加拿大	1.38	12	英国	1.14	17	法国	0.74
3	比利时	1.62	8	新加坡	1.37	13	意大利	1.00	18	印度	0.69
4	丹麦	1.57	9	澳大利亚	1.30	14	西班牙	0.84	19	巴西	0.46
5	中国	1.53	10	瑞典	1.21	15	德国	0.83			


注:μR－被引率均值

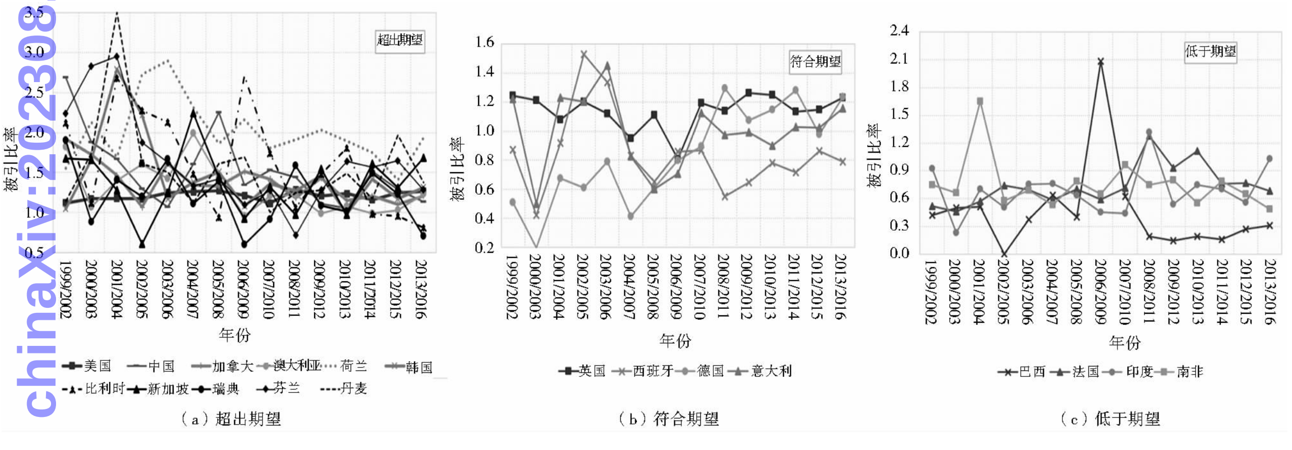


图 6　归一化被引比率期望子群

(2)符合期望(0.8≤RoC<1.2):英国、意大利、西班牙、德国四国的 RoC 均值符合期望,其中英国 RoC 均值为 1.14,该国在新的科学技术迅猛发展与研究范式转型的大背景下仍然保持了较高质量的学术论文产出水平。

(3)低于期望(RoC<0.8):巴西、法国、印度、南非四国的 RoC 均值低于期望,结合上文分析,四国在学术论文产出数量和质量两个维度上都处于追赶地位。

5　学术论文产出的主题知识图谱分析

关键词是对作者对文章研究内容的精炼,一定程度上能够反映文章的研究主题,是文献计量学研究中的重要指标[12]。基于关键词共现矩阵构建主题知识图谱,可以对学科领域内的研究内容进行可视化解读,从而更好把握学科的知识发展脉络与研究前沿[13]。本文从时间窗口和国家两个维度对图情领域的主题知识图谱展开分析。

5.1　不同时间窗口的主题知识图谱演化

5.1.1　主题知识网络图谱绘制　以数据集 D1 为来源数据,将近 15 年数据划分为时间间隔相等的三个时间窗口 T1(2002－2006 年)、T2(2007－2011 年)、T3(2012－2016 年),对 T1、T2、T3 时间窗口下的图情领域研究主题演变展开分析。本文采用 J. C. Donohue 的高频词临界算法(公式 1)对清洗后的高频关键词数量加以计算。T1 时期计算得到高频关键词 120 个,总频次 4 515 次,占 9 443 个关键词总频次的 56.15%;T2

时期计算得到高频关键词 172 个,总频次 10 583 次,占 19 460 个关键词总频次的 59.03%;T3 时期计算得到高频关键词 219 个,总频次 17 902 次,占 31 757 个关键词总频次的 63.23%。采用数据分析软件 Thomson Data Analyzer 分别构建高频关键词共现矩阵,并利用可视化软件 NetDraw 绘制各时间窗口的主题知识图谱(见图 7)。由图 7 可知,三个时间窗口下的知识图谱整体网络结构较为紧密,且网络密度不断增加,高频关键词节点之间的共现行为趋向频繁。

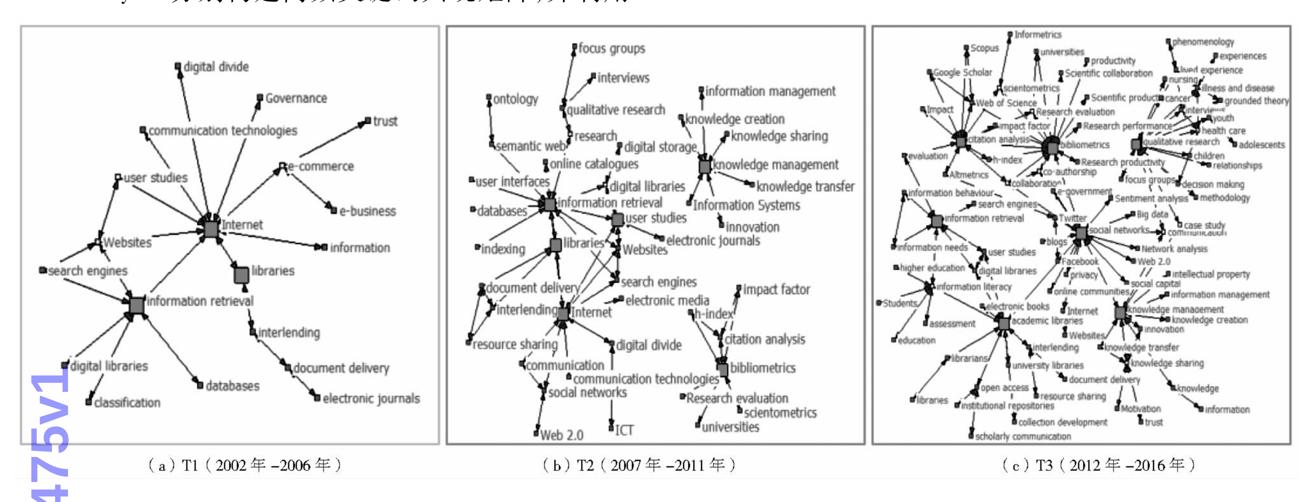


图 7 三个时间窗口的主题知识图谱(共现频次≥10)

5.1.2 主题范畴分析 基于主题知识图谱,图情领域的研究主题可以概括为图书馆、信息检索、互联网、知识管理、社交网络、计量学、用户研究、研究方法 8 个主题范畴。具体地,对三个时间窗口下的研究主题范畴详细分析如下。

T1 时期图情领域的主题范畴包括“互联网”“图书馆”以及“信息检索”。其中,“互联网”与电子商务、信息传播技术、数字鸿沟等关键词紧密关联;“图书馆”与馆际互借、文献传递、电子期刊等关键词紧密关联;“信息检索”与搜索引擎、分类法、数据库等关键词紧密关联。

T2 时期图情领域的主题范畴包括“互联网”“信息检索”“图书馆”“用户研究”“知识管理”以及“文献计量学”。其中,“互联网”从 Web1.0 转向 Web2.0 情境,“信息检索”与语义网、电子图书馆、电子目录的关联趋向紧密;“用户研究”主要与信息检索、互联网等主题相关;“知识管理”与知识分享、知识转移、知识创造、创新等关键词紧密关联;“文献计量学”与引文分析、影响因子、研究评价、科学计量学、高校等关键词紧密关联。

T3 时期图情领域的主题范畴包括“信息检索”“学术图书馆”“知识管理”“文献计量学”“社交网络”以及“定性研究方法”。其中,“信息检索”的研究重心转向信息需求、信息行为、信息素养;“知识管理”对知识产权问题日益重视,社交网络环境下的知识管理研究成为新的研究方向,如基于社交网络信息的自然灾害预测研究[14];“学术图书馆”与开放获取、电子图书馆、社交网络、信息素养、图书馆员等关键词紧密关联;“文献计量学”主要用于测度科研表现、科学生产力、科研合作以及研究评价;“社交网络”的研究对象主要包括 Twitter、Blogs、Facebook,与情感分析、网络分析、信息交流、大数据、隐私、网络社区等关键词紧密关联;“定性研究方法”与医疗健康主题的共现最为频繁,最常使用的研究方法包括访谈法、案例分析、焦点小组、扎根理论。

5.1.3 高被引期刊论文分析 对不同时间窗口、各主题范畴内的学术论文产出进行内容分析,近 15 年内 ESI 高被引期刊论文及其作者等信息如表 4 所示:

表 4 图情领域高被引论文

排序	作者	论文名称	来源期刊
1	N. B. Ellison[15]	社会网络站点:定义、历史和学术研究	JCMC
2	S. Petter[16]	信息系统研究中的形成性构念的确定	MISQ
3	V. Venkatesh[17]	消费者信息技术的接受和使用:信息技术接受使用整合模型	MISQ
4	N. D. Liben[18]	社会网络的链接预测问题	JASIST
5	M. Wetzels[19]	用 PLS 路径模型评价层次结构模型:准则与实证	MISQ
6	H. Chen[20]	商业智能与分析:从大数据到大冲击	MISQ
7	K. Peffers[21]	用于信息系统研究的一种设计科学研究方法论	JMIS
8	S. Valenzuela[22]	社交网站中是否有社会资本? Facebook 使用与大学生满意度、信任、参与	JCMC
9	B. J. Jansen[23]	Twitter 的力量:作为电子口碑的推文	JASIST
10	C. L. Hsu[24]	博客的接受和使用:技术接受、社会影响和知识分享动机的作用	I&M

由表4可知,近十五年图情领域的ESI高被引论文主题集中在社交网络、用户行为、大数据、信息系统等领域,来源期刊以MISQ(*MIS Quarterly*)、JCMC(*Journal of computer mediated communication*)、JASIST(*Journal of the Association for Information Science and Technology*)为主。

5.2 国家－主题二模知识图谱分析

5.2.1 二模知识图谱绘制　二模网络用于分析两种不同类型模态节点之间的网络关系,可细分为二元双模网络与从属网络[25]。本文构建的二模网络的两种节点分别为国家及关键词,属于从属网络。通过构建国家－主题二模网络并绘制二模知识图谱,可以直观揭示各国在图情领域的研究主题分布及不同国家之间以研究主题为中介所产生的关联。

数据集D2中在统计各国的发文量时考虑了所有作者的学术贡献,因此在统计各国发文量时存在重复统计,但是好处是可以更加全面地考察所有排序作者所在国家对不同主题领域的学术贡献。本文以数据集D2作为国家主题分析的数据来源,选取频次在100及以上的关键词共114个,构建19行×114列的二模矩阵。由于美国等国家的学术论文产出丰富,国家－主题之间的关系网络结构较为复杂,因此在绘制知识图谱时,本文对二模矩阵数据进行了精简:对各国的国家－主题关联强度分别计算四分位数,保留排序在前25%的关键词;其次,删除关联强度低于25次的所有关键词,最终得到包含19个国家节点、65个关键词节点的二模矩阵,将其转换为二值矩阵导入UCinet绘制二模知识图谱如图8所示:

图8　国家－主题二模知识图谱

5.2.2 二模网络中心度分析　二模网络中国家节点与关键词节点的中心度分析结果见表5。结合表5、图8可知,社交网络、信息检索、知识管理是近15年图情领域的研究热点,研究力量最为强大;英国、西班牙、中国、美国的研究主题最为广泛,除了中心度值较高的若干关键词以外还形成了自身研究特色。英国在文献传递、公共图书馆、电子图书等图书馆学主题范畴的成果丰硕;西班牙在网络计量、语义网、研究评价等计量学主题范畴的科研表现最为突出;中国在文本挖掘、数据挖掘、H指数等计量学主题范畴以及用户满意、技术接受模型等用户研究主题范畴的研究成果非常醒目;美国在数字鸿沟、数字图书馆等图书馆学主题范畴、信任、隐私等用户研究主题范畴、知识共享、决策支持等知识管理主题范畴都进行了非常深入的研究。

表5　二模网络中心度

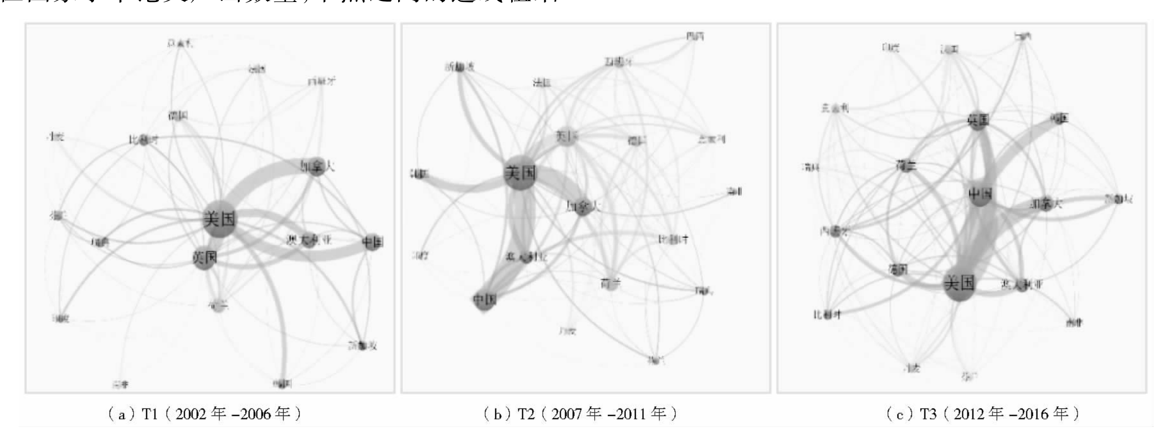
序号	国家	中心度	关键词	中心度
1	英国	28	社交网络	11
2	西班牙	28	信息检索	11
3	中国	27	知识管理	11
4	美国	27	文献计量学	10
5	加拿大	14	互联网	6

6　重要国家学术合作影响因素分析

6.1　基于国家合作知识图谱的影响因素发现

随着科学技术的发展和学科国际合作交流程度的不断加深,图情领域的国际学术合作格局也在发生相

应改变。本文以 19 个重要国家为研究对象，构建三个时间窗口下的重要国家学术合作图谱，以期深入挖掘国家之间的学术合作行为机理。如图 9 所示，节点大小表征国家学术论文产出数量，节点之间的连线粗细和距离表征国家间学术合作的密切程度，节点颜色表征基于 VOSviewer 聚类算法被划分为不同学术合作类型的国家。



（a）T1（2002 年 –2006 年）（b）T2（2007 年 –2011 年）（c）T3（2012 年 –2016 年）

图 9 三个时间窗口的学术合作图谱

由图 9 可知，从 2002 年至今，重要国家之间的学术合作趋向频繁，特别是在 T3 时期，重要国家之间的学术合作网络密度高达 0.93，多数国家间通过文献合著方式产生了学术交流合作行为。对比分析 T1、T2、T3 的国家学术合作图谱，美国、英国、中国、加拿大中心度排名前四的国家始终处于合作中心，与其他国家都形成了较为巩固的合作关系。对比之下，其他 15 个重要国家在不同时期内的学术合作对象相对固定，合作对象数量稳中有升。本文尝试将学术合作阈值设定为 10 次后发现，各国之间的合作对象呈现出明显的地区分布特征，瑞典的学术合作对象仅为同样位于欧洲的荷兰、德国、芬兰三个国家，新加坡的合作对象仅为同样位于亚洲的韩国。Acosta 等[26]、Parreira 等[27]学者研究同样指出，空间距离是影响学者学术合作的重要因素，随着空间距离的增加，学术合作的成本明显提升。

6.2 学术合作影响因素模型的实证分析

上文分析表明空间距离是影响国际学术合作的重要因素，同时，诸如美国等已经具备学术能力累积优势的国家在学术合作中能够获取更多的优势资源，间接消减了空间距离带来的消极影响。鉴于此，本文参考张萃、欧阳冬平[28]的学术合作引力模型并构建图情领域学术合作影响因素计量模型（公式 3）。式中，$H_{ij,t}$是 t 时期 i、j 两国的学术合作学术论文产出数量；$P1_{it}$和 $P2_{jt}$分别是 t 时期 i 国自身和学术合作对象 j 国的学术论文产出总量，用于表征两国的学术能力；D_{ij}是 i、j 两国的空间距离；$\varepsilon_{ij,t}$是误差项。

$$H_{ij,t} = \alpha_0 + \alpha_1 \ln P1_{i,t} + \alpha_2 \ln P2_{j,t} + \alpha_3 \ln D_{ij} + \varepsilon_{ij,t} \quad (3)$$

根据 19 个国家首都的经纬度计算得到各国两两之间的距离，以此表征国家空间距离。19 个国家中包含 4 个亚洲国家、10 个欧洲国家、2 个北美洲国家以及南美洲、大洋洲、非洲国家各 1 个。分别以亚洲、欧洲、北美洲三个地区和 T1、T2、T3 三个时间窗口为计量单元，对近 15 年内世界范围内各大洲国家的学术合作学术论文产出模型和不同时间窗口的 19 个国家学术合作学术论文产出模型分别进行实证分析。由于 $H_{ij,t}$是数值型变量，故而选用负二项回归模型进行回归分析较为适宜。应用数据分析软件 Stata14.0 得到实证结果如表 6 所示：

表 6 学术合作学术论文产出模型实证分析结果

变量	整体	地区			时间窗口		
		亚洲	欧洲	北美洲	T1	T2	T3
lnP1	0.89 ***(0.04)	0.97 ***(0.14)	0.76 ***(0.07)	0.89 ***(0.08)	0.79 ***(0.05)	0.91 ***(0.05)	0.92 ***(0.04)
lnP2	0.89 ***(0.04)	1.14 ***(0.12)	0.79 ***(0.05)	1.06 ***(0.11)	0.79 ***(0.05)	0.92 ***(0.05)	0.92 ***(0.04)
lnD	-0.38 ***(0.04)	-0.84 **(0.28)	-0.39 ***(0.04)	0.02(0.11)	-0.34 ***(0.05)	-0.35 ***(0.04)	-0.36 ***(0.04)
Constant	-7.66 ***(0.46)	-6.83 ***(2.04)	-5.93 ***(0.62)	-11.36 ***(1.53)	-2.68 ***(0.78)	-4.08 ***(0.71)	-3.99 ***(0.61)

注：***、**分别表示变量系数在 0.001、0.01 的概率水平下显著，系数下方括号内数值为标准误

实证分析结果表明,无论是国家自身的科研实力水平还是学术合作对象的学术论文产出能力都对两国之间学术合作与学术交流存在显著正向影响,而两国之间的空间距离则对学术合作具有明显的负面影响,但学术能力的影响力比空间距离更大,因此提升自身在图情领域的学术论文产出质量是打破空间隔离、促进国际交流的有效方式。

(1)学术能力对学术合作的影响:亚洲国家与其他国家的学术合作受学术能力的影响更大;在三大洲各国中,相较于自身的学术能力,合作对象学术能力对学术合作行为的影响更为突出。随着时间的推移,学术能力对学术合作的影响程度逐渐加深。

(2)空间距离对学术合作的影响:从地区维度来看,亚洲国家(中国、韩国、新加坡、印度)的学术合作受空间距离的影响最大,地区内部展开地学术合作活动相对频繁,但与欧洲、北美洲等其他各大洲国家的学术交流程度还有待提升;欧洲国家(英国、西班牙、澳大利亚等)的学术合作也呈现一定的地区指向性;北美洲国家(美国、加拿大)的学术合作则不受空间距离因素的阻碍,两国在图情领域的学术研究具有更好的国际视野。从时间维度来看,随着时间的推移,各国之间的学术交流活动的地区化趋势反而有所增强,与科学交流日趋国际化的固有认知相左,这或与地区内部频繁的政治、经济、文化合作存在一定关联,区域性的国际组织如欧盟、亚太经合组织对地区内部国家之间的学术文化交流提供了一定的便利性条件。

7 结论与讨论

本文以图情领域 19 个重要国家为研究对象,搜集这些国家在近 15 年内的学术论文产出与引证数据,构建共词矩阵并绘制科学知识图谱,从时序视域对这些国家在图情领域的学术论文产出贡献、引证贡献、研究主题贡献分别进行了对比分析,并对各国之家的学术交流合作的影响机理加以揭示。无论从学术论文产出总量还是引证数量来看,美国在图情领域的研究实力雄厚,学术贡献最为突出;中国在图情领域的学术论文产出数量增速明显,在近五年内其学术论文产出排名已稳居第二;英国、加拿大、西班牙、澳大利亚等国在图情领域各维度的学术贡献也位居前列;同时,荷兰、芬兰、比利时等 11 国在归一化被引比率指标上超出期望。从主题维度来看,近 15 年内图情领域的研究主题主要集中在图书馆、信息检索、互联网、知识管理、社交网络、计量学、用户研究、研究方法 8 个主题范畴,英国、西班牙、中国、美国的研究主题最为广泛。从学术合作的影响因素来看,空间距离和学术能力是影响国家间学术合作行为的重要因素,随着时间的推移两个指标的影响力都在逐渐增加。在三大洲各国中,中国、韩国、新加坡、印度 4 个亚洲国家的学术合作受空间距离、学术能力两个指标的影响更大。

本文还存在以下研究局限:本文从学术论文层面较为全面地揭示了国际图书情报领域各国的学术贡献和学术合作的影响因素,但是对各国在国际会议、重要学术机构、重要期刊评审中的参与程度和贡献没有进行考察,各国学术研究的社会影响力也有待检验,这将是本研究的后续努力方向。

参考文献:

[1] 段先娥, 张晓娟. 中国图书情报领域国际发文状况分析 (2011-2015)[J]. 图书馆论坛, 2017, 37(5): 64-73.

[2] 施萧萧, 张庆普. 基于共词分析的国外颠覆性创新研究现状及发展趋势[J]. 情报学报, 2017, 36(7): 748-759.

[3] 李信, 李旭晖, 陆伟, 等. 大数据驱动下的图书情报学科热点领域挖掘——面向 WOS 题录数据的实证视角[J]. 图书馆论坛, 2017, 37(4): 49-57.

[4] 谭春辉, 曾娟, 邱均平. 基于 CSSCI 的"十二五"时期国内情报学研究态势分析[J]. 情报学报, 2017, 36(7): 734-747.

[5] 赵蓉英, 魏绪秋. 基于比较静态分析的国际合作与交流特征演化研究——以我国图书情报学为例[J]. 图书馆, 2017 (8): 27-31.

[6] 童寿传, 李江, 李东. 我国图书情报学国际化发展的现状与趋势[J]. 图书情报知识, 2017 (6): 24-34.

[7] 侯剑华, 杨秀财, 周莉娟. 国际图书情报领域研究的前沿主题及其演化趋势分析[J]. 图书情报工作, 2016, 60(13): 82-90.

[8] DONOHUE J C. Understanding scientific literatures: a bibliometric approach[M]. Cambridge: MIT Press, 1973:129.

[9] BORNMANN L, WAGNER C, LEYDESDORFF L. The geography of references in elite articles: which countries contribute to the archives of knowledge? [J]. PloS one, 2018, 13(3): e0194805.

[10] VAN ECK N J, WALTMAN L. Software survey: vosviewer, a computer program for bibliometric mapping[J]. Scientometrics, 2010, 84(2): 523-538.

[11] 何汶, 刘颖, 杨红梅. 图情领域高被引论文国际合著及贡献研究——基于文献计量学角度[J]. 图书情报知识, 2017(2):83-93.

[12] 柯平, 宫平. 数字人文研究演化路径与热点领域分析[J]. 中国图书馆学报, 2016, 42(6): 13-30.

[13] 胡泽文, 廉同辉, 曹玲. 网络资源聚合与导航研究的科学知识图谱分析[J]. 情报科学, 2017, 35(7): 9-15.

[14] HERFORT B, BRENNING A, ZIPF A. A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management[J]. International jour-

nal of geographical information science, 2015, 29(4): 667 – 689.
[15] ELLSION N B. Social network sites: definition, history, and scholarship[J]. Journal of computer mediated communication, 2007, 13(1): 210 – 230.
[16] PETTER S, STRAUB D, RAI A. Specifying formative constructs in information systems research[J]. MIS quarterly, 2007, 31(4): 623 – 656.
[17] VENKATESH V, THONG J Y L, XU X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology[J]. MIS quarterly, 2012, 36(1): 157 – 178.
[18] LIBEN NOWELL D, KLEINBERG J. The link prediction problem for social networks[J]. Journal of the Association for Information Science and Technology, 2007, 58(7): 1019 – 1031.
[19] WETZELS M, ODEKERKEN SCHRODER G, OPPEN C V. Using PLS path modeling for assessing hierarchical construct models: guidelines and empirical illustration[J]. MIS quarterly, 2009, 33(1): 177 – 195.
[20] CHEN H, CHIANG R H L, STOREY V C. Business intelligence and analytics: from big data to big impact[J]. MIS quarterly, 2012, 36(4): 1165 – 1188.
[21] PEFFERS K, TUUNANEN T, ROTHENBERGER M A, et al. A design science research methodology for information systems research[J]. Journal of management information systems, 2007, 24(3): 45 – 77.
[22] VALENZUELA S, PARK N, KEE K F. Is there social capital in a social network site? Facebook use and college students' life satisfaction, trust, and participation[J]. Journal of computer-mediated communication, 2009, 14(4): 875 – 901.
[23] JANSEN B J, ZHANG M, SOBEL K, et al. Twitter power: tweets as electronic word of mouth[J]. Journal of the Association for Information Science and Technology, 2009, 60(11): 2169 – 2188.
[24] HSU C L, LIN J C C. Acceptance of blog usage: the roles of technology acceptance, social influence and knowledge sharing motivation[J]. Information & management, 2008, 45(1): 65 – 74.
[25] BORGATTI S P, MEHRA A, BRASS D J, et al. Network analysis in the social sciences[J]. Science, 2009, 324(5926): 892 – 895.
[26] ACOSTA M, CORONADO D, FERRANDIZ E, et al. The geography of university scientific production in Europe: an exploration in the field of food science and technology[J]. Scientometrics, 2017, 112(1): 215 – 240.
[27] PARREIRA M R, MACHADO K B, LOGARES R, et al. The roles of geographic distance and socioeconomic factors on international collaboration among ecologists[J]. Scientometrics, 2017, 113(3): 1539 – 1550.
[28] 张萃, 欧阳冬平. "一带一路"战略下中国国际学术合作影响因素研究——基于 Web of Science 数据库中外合作科研论文的实证分析[J]. 国际贸易问题, 2017 (4): 74 – 82.

作者贡献说明:
刘虹:框架设计、论文撰写与修订;
李煜:数据处理。

Research on Academic Contributions of Major Countries of Library and Information Science Based on Time Series Perspective

Liu Hong　Li Yu
School of Information Management, Nanjing University, Nanjing 210023

Abstract: [**Purpose/significance**] The article reveals the academic contributions and evolution features of major countries of library and information science since 2002 on the dimensions of outputs of papers, citations and research topics, in the hope of providing a reference for the subject construction and related researchers tracking research topics and frontiers. [**Method/process**] Based on 86 core journals in the last 15 years of library and information science in Web of Science, the article used research methods including bibliometrics, scientific knowledge maps and econometric models, utilized data analysis tools including TDA, Ucinet and VOSviewer. The article gave an in-depth analysis of the academic contribution of major countries of library and information science from the following perspectives: academic papers amount, citation features and theme features. Meanwhile, the article built a metrological model to examine the impact of spatial distance and academic ability on academic cooperation among major countries. [**Result/conclusion**] The United States has a strong scientific research strength in the field of library and information science, China has the fastest growth of academic contribution, Netherlands, Finland and Belgium all excel in quoted rate indicators. The research topics mainly focus on the eight thematic areas of libraries, information retrieval, internet, knowledge management, social networks, metrology, user study and research methods. Britain, Spain, China and the United States have the most extensive research topics. Spatial distance and academic capabilities have a significant impact on academic cooperation among countries.

Keywords: library and information science　scientific collaboration　informetrics　knowledge map　visualization